DSIC WORKING GROUP ISSUES

I. **Scope/objectives:**

A. Make recommendations to caBIG Oversight Board.concerning data sharing and protection of intellectual capital created by caBIG participants; develop standards, draft policy documents and white papers as needed.

i. Apart from traditional notions of intellectual property, what do we mean when we use the term "intellectual capital" and what are we trying to protect?

ii. How does traditional intellectual property (copyright, patent rights, trademarks, trade secrets) fit into caBIG objectives and needs of stakeholders?

B. Define scope and goals of cross initiative activities within caBIG
   i. solicit input to ensure that other issues are addressed
   ii. provide expert guidance regarding specific areas of concern raised within caBIG Workspaces and individual Project Teams associated with these issues

C. Stakeholders:

i. Patients

ii. Government
   1. immediately as funder of research
   2. ultimately as purchaser of healthcare through Medicare/Medicaid, looking for cost effective solutions

iii. Academic institutions
   1. scientists/developers
   2. administration (tech transfer & sponsored research offices)

iv. Industry
   1. tool developers/providers
   2. pharma/biotech developers of drugs & diagnostics

v. Other??

D. Solicit additional legal/regulatory expertise as needed:

i. An organization representing IRBs (e.g., Public Responsibility in Medicine and Research (PRIM&R))

ii. American Bar Association (ABA) and/or American Intellectual Property Law Association (AIPLA)

iii. FDA and/or NCI/CTEP

II. Software tools

A. Copyright vs. open source

i. Discuss meaning of open source software as defined in subcontracts between master contractor and cancer centers

ii. What are the implications of the caBIG definition of open source for work with industry partners?
   1. Tool companies working with cancer center to develop software
   2. Recipients of software developed by cancer centers

iii. Presumably cancer center contracts administrators will review institutional policies to ensure that standard copyright policies (and individual employment contracts of cancer center employees working caBIG) permit open source or will obtain necessary waivers.

B. Patent rights

i. Incompatible with open source?
   1. If not, will cancer center investigators develop patentable bioinformatics software tools?
   2. If so, these institutions will have rights and obligations under the Bayh-Dole Act; federal regulations require that certain procedures be carried out to modify these rights, e.g., disclaiming any rights to file patents, etc..

III. Sharing data

A. Definitions

i. Raw data (clinical information and specimens)

ii. Derivative data – clinomic,genomic, proteomic (e.g., output of microarray experiments, clinical trials, etc. – can we clearly delineate all possible derivative outputs?

iii. Apart from restrictions described below, what is meant by an "open data" policy? Viral vs. non-viral?

B. Examples of other approaches (funded in whole or in part by NIH):

i. AP4
ii. EDRN
iii. Penn Cancer Alliance

iv. HapMap
v. Consortium for Functional Glycomics
vi. Alliance for Cellular Signaling
vii. Human Genome Project
viii. SNP Consortium
ix. Protein Structure Initiative
x. Integrative Cancer Biology Program
xi. other

C. Potential restrictions on access to data

i. Patient-related issues

1. Human subjects rights -- three issues concerning patient consent that must be addressed:
    a. data that can be shared is dependent upon the consent given by the patient;
    b. the scope of patient consent is haphazard; and
    c. the wording of IRB consent forms needs to be standardized.

2. Privacy rights (HIPAA) -- it is not possible to completely de-identify genomic data. Existing privacy authorization forms can be used a baseline, but the scope of these consents seem almost random.

3. Possible solution: Recommend ways to standardize informed consent documents/processes and HIPAA privacy authorizations.

ii. Scientists' unwillingness to share –

1. risk of having data collected by primary investigator interpreted out of context by secondary investigator (e.g., partial datasets made available before conclusion of long term studies) –

2. desire to fully analyze unpublished data before giving others a chance to evaluate -- primary investigator gets to publish first;

3. concerns re: technical accuracy – will data be sent and maintained accurately in storage?

4. Possible solutions: Develop mechanisms for accrediting primary investigators for development of data in subsequent, follow-on research; create incentives for collaborations between primary and secondary investigators.

iii. Institutional concerns driving restrictions on access to data

1. risk of noncompliance with HIPAA

2. Assure priority for filing patent applications; protect licensing relationships and future opportunities

3. Protect industry proprietary information (trade secrets, regulatory filings,?)

    a. How to ensure access to discoveries made with an individual industry collaborator's information (e.g., new targets/markers achieved through use of proprietary compounds) by other industry participants (i.e., competitors)? Will such collaborators assert reach-through rights?

4. Possible solutions: solicit input from champions within institutional technology transfer offices and pharma/biotech research communities to develop realistic standards for data sharing in academic-industry collaborations.

D. Immediate opportunities for data sharing (with follow-on issues):

i. Published data and supporting data publications

1. Clinical trials (CT)–

a. Share all institutional trial information -- is there a prototypical system that could be leveraged?
b. Start with all open trials and include Pharma trials
    i. Need input from industry & FDA/CTEP to understand perceived legal barriers to sharing data & protocols
c. Variable views should be available:
    i. Public view: make available in first release
    ii. Participating patients: make available in future releases; need to determine what is required for this view
    iii. Referring physicians: make available in future releases; need to determine what is required for this view
d. Patient Privacy – see discussion above; need to ensure that private information not disclosed

-

2. **Other? --** Microarray experiments? Other data coming out of ICR workspace?

ii. Specimen and tissue resources

1. Develop a virtual repository of de-identified specimen and tissue resources using the EDRN model
    a. setting up the systems and establishing the data exchanges is the most difficult

b. mapping data elements, which are not consistent, is difficult
c. need consistent annotations across all sites
d. research evaluation management process may be required
e. need standardized material transfer/licensing agreeements (see model agreements at www.pcabc.upmc.edu) and institutional commitment to streamline the transfer process, especially in terms of access for industry (pharma/ biotech)

E. Longer term opportunities

i. Develop secure and technically accurate methods of sharing pre-publication on a tiered basis, which will protect novel IP and assure that appropriate investigators have priority for publication purposes

1. Relevant to which workspaces -- ICR? Other?

ii. Develop standards for sharing data that are acceptable to industry and academic researchers

1. educate biotech/pharma as to benefits of open data policy and lost opportunities resulting from failure to share data, including negative data
2. need to understand biotech/pharma's perceived risks associated with sharing data (protection of proprietary information from competitors).